(54) Title: DECODING OF ARRAY SENSORS WITH MICROSPHERES

(57) Abstract

The invention relates to compositions and methods for decoding microsphere array sensors.

DECODING OF ARRAY SENSORS WITH MICROSPHERES

This application is a continuing application of U.S.S.N.s 09/189,543, filed November 10, 1998, and of
60/090,473, filed June 24, 1998.

FIELD OF THE INVENTION

The invention relates to compositions and methods for decoding microsphere array sensors.

BACKGROUND OF THE INVENTION

There are a number of assays and sensors for the detection of the presence and/or concentration of
specific substances in fluids and gases. Many of these rely on specific ligand/antiligand reactions as
the mechanism of detection. That is, pairs of substances (i.e. the binding pairs or ligand/antiligands)
are known to bind to each other, while binding little or not at all to other substances. This has been the
focus of a number of techniques that utilize these binding pairs for the detection of the complexes.
These generally are done by labeling one component of the complex in some way, so as to make the
entire complex detectable, using, for example, radioisotopes, fluorescent and other optically active
molecules, enzymes, etc.

Of particular use in these sensors are detection mechanisms utilizing luminescence. Recently, the
use of optical fibers and optical fiber strands in combination with light absorbing dyes for chemical
analytical determinations has undergone rapid development, particularly within the last decade. The
use of optical fibers for such purposes and techniques is described by Milanovich et al., "Novel Optical
Fiber Techniques For Medical Application", Proceedings of the SPIE 28th Annual International
Technical Symposium On Optics and Electro-Optics, Volume 494, 1980; Seitz, W.R., "Chemical
Sensors Based On Immobilized Indicators and Fiber Optics" in C.R.C. Critical Reviews In Analytical
Chemistry, Vol. 19, 1988, pp. 135-173; Wolfbeis, O.S., "Fiber Optical Fluorosensors In Analytical

Chemistry" in *Molecular Luminescence Spectroscopy, Methods and Applications* (S. G. Schulman, editor), Wiley & Sons, New York (1988); Angel, S.M., *Spectroscopy 2* (4):38 (1987); Walt, et al., "Chemical Sensors and Microinstrumentation", *ACS Symposium Series*, Vol. 403, 1989, p. 252, and Wolfbeis, O.S., *Fiber Optic Chemical Sensors*, Ed. CRC Press, Boca Raton, FL, 1991, 2nd Volume.

5

When using an optical fiber in an *in vitro/in vivo* sensor, one or more light absorbing dyes are located near its distal end. Typically, light from an appropriate source is used to illuminate the dyes through the fiber's proximal end. The light propagates along the length of the optical fiber; and a portion of this propagated light exits the distal end and is absorbed by the dyes. The light absorbing dye may or may

10   not be immobilized; may or may not be directly attached to the optical fiber itself; may or may not be suspended in a fluid sample containing one or more analytes of interest; and may or may not be retainable for subsequent use in a second optical determination.

Once the light has been absorbed by the dye, some light of varying wavelength and intensity returns,

15   conveyed through either the same fiber or collection fiber(s) to a detection system where it is observed and measured. The interactions between the light conveyed by the optical fiber and the properties of the light absorbing dye provide an optical basis for both qualitative and quantitative determinations.

Of the many different classes of light absorbing dyes which conventionally are employed with bundles

20   of fiber strands and optical fibers for different analytical purposes are those more common compositions that emit light after absorption termed "fluorophores" and those which absorb light and internally convert the absorbed light to heat, rather than emit it as light, termed "chromophores."

Fluorescence is a physical phenomenon based upon the ability of some molecules to absorb light

25   (photons) at specified wavelengths and then emit light of a longer wavelength and at a lower energy. Substances able to fluoresce share a number of common characteristics: the ability to absorb light energy at one wavelength λ_{ab}; reach an excited energy state; and subsequently emit light at another light wavelength, λ_{em}. The absorption and fluorescence emission spectra are individual for each fluorophore and are often graphically represented as two separate curves that are slightly overlapping.

30   The same fluorescence emission spectrum is generally observed irrespective of the wavelength of the exciting light and, accordingly, the wavelength and energy of the exciting light may be varied within limits; but the light emitted by the fluorophore will always provide the same emission spectrum. Finally, the strength of the fluorescence signal may be measured as the quantum yield of light emitted. The fluorescence quantum yield is the ratio of the number of photons emitted in comparison to the number

35   of photons initially absorbed by the fluorophore. For more detailed information regarding each of these characteristics, the following references are recommended: Lakowicz, J. R., *Principles of Fluorescence Spectroscopy*, Plenum Press, New York, 1983; Freifelder, D., Physical Biochemistry,

second edition, W. H. Freeman and Company, New York, 1982; "Molecular Luminescence

Spectroscopy Methods and Applications: Part I" (S.G. Schulman, editor) in *Chemical Analysis*, vol. 77,

Wiley & Sons, Inc., 1985; *The Theory of Luminescence*, Stepanov and Gribkovskii, Iliffe Books, Ltd.,

London, 1968.

5

In comparison, substances which absorb light and do not fluoresce usually convert the light into heat

or kinetic energy. The ability to internally convert the absorbed light identifies the dye as a

"chromophore." Dyes which absorb light energy as chromophores do so at individual wavelengths of

energy and are characterized by a distinctive molar absorption coefficient at that wavelength.

10   Chemical analysis employing fiber optic strands and absorption spectroscopy using visible and

ultraviolet light wavelengths in combination with the absorption coefficient allow for the determination

of concentration for specific analyses of interest by spectral measurement. The most common use of

absorbance measurement via optical fibers is to determine concentration which is calculated in

accordance with Beers' law; accordingly, at a single absorbance wavelength, the greater the quantity

15   of the composition which absorbs light energy at a given wavelength, the greater the optical density for

the sample. In this way, the total quantity of light absorbed directly correlates with the quantity of the

composition in the sample.

Many of the recent improvements employing optical fiber sensors in both qualitative and quantitative

20   analytical determinations concern the desirability of depositing and/or immobilizing various light

absorbing dyes at the distal end of the optical fiber. In this manner, a variety of different optical fiber

chemical sensors and methods have been reported for specific analytical determinations and

applications such as pH measurement, oxygen detection, and carbon dioxide analyses. These

developments are exemplified by the following publications: Freeman, et al., *Anal Chem.* 53:98 (1983);

25   Lippitsch et al., *Anal. Chem. Acta.* 205:1, (1988); Wolfbeis et al., *Anal. Chem.* 60:2028 (1988); Jordan,

et al., *Anal. Chem.* 59:437 (1987); Lubbers et al., *Sens. Actuators* 1983; Munkholm et al., *Talanta*

35:109 (1988); Munkholm et al., *Anal. Chem.* 58:1427 (1986); Seitz, W. R., *Anal. Chem.* 56:16A-34A

(1984); Peterson, et al., *Anal. Chem.* 52:864 (1980); Saari, et al., *Anal. Chem.* 54:821 (1982); Saari, et

al., *Anal. Chem.* 55:667 (1983); Zhujun et al., *Anal. Chem. Acta.* 160:47 (1984); Schwab, et al., *Anal.*

30   *Chem.* 56:2199 (1984); Wolfbeis, O.S., "Fiber Optic Chemical Sensors", *Ed. CRC Press, Boca Raton,*

FL, 1991, 2nd Volume; and Pantano, P., Walt, D.R., *Anal. Chem.*, 481A-487A, Vol. 67, (1995).

More recently, fiber optic sensors have been constructed that permit the use of multiple dyes with a

single, discrete fiber optic bundle. U.S. Pat. Nos. 5,244,636 and 5,250,264 to Walt, *et al.* disclose

35   systems for affixing multiple, different dyes on the distal end of the bundle, the teachings of each of

these patents being incorporated herein by this reference. The disclosed configurations enable

separate optical fibers of the bundle to optically access individual dyes. This avoids the problem of

deconvolving the separate signals in the returning light from each dye, which arises when the signals from two or more dyes are combined, each dye being sensitive to a different analyte, and there is significant overlap in the dyes' emission spectra.

5   U.S.S.N.s 08/818,199 and 09/151,877 describe array compositions that utilize microspheres or beads on a surface of a substrate, for example on a terminal end of a fiber optic bundle, with each individual fiber comprising a bead containing an optical signature. Since the beads go down randomly, a unique optical signature is needed to "decode" the array; i.e. after the array is made, a correlation of the location of an individual site on the array with the bead or bioactive agent at that particular site can be 10   made. This means that the beads may be randomly distributed on the array, a fast and inexpensive process as compared to either the *in situ* synthesis or ˜˜otting techniques of the prior art. Once the array is loaded with the beads, the array can be decoded, or can be used, with full or partial decoding occurring after testing, as is more fully outlined below.

15   One drawback with the previous system is that it requires a set of unique optical signatures. While large sets of such signatures are available, for example by using different ratios of different dyes, it would be preferable to use decoding systems that do not rely on the use of sets of optical signatures. Accordingly, it is an object of the invention to provide methods to allow decoding of bead arrays without relying solely on unique optical signatures.
20

SUMMARY OF THE INVENTION

In accordance with the above objects, the present invention provides array compositions comprising a substrate with a surface comprising discrete sites. The composition further comprises a population of 25   microspheres comprising at least a first and a second subpopulation; each subpopulation comprises a bioactive agent; and an identifier binding ligand that will bind a decoder binding ligand such that the identity of the bioactive agent can be elucidated. The microspheres are distributed on the surface.

In a further aspect, the present invention provides array compositions comprising a substrate with a 30   surface comprising discrete sites, and a population of microspheres comprising at least a first and a second subpopulati˜˜ Each subpopulation comprises a bioactive agent and does not comprise an optical signature.

In an additional aspect, the present invention provides methods of making an array composition as 35   outlined above. The methods comprise forming a surface comprising individual sites on a substrate and distributing microspheres on said surface such that said individual sites contain microspheres. The microspheres comprise at least a first and a second subpopulation each comprising a bioactive

agent and do not comprise an optical signature.

In a further aspect, the invention provides methods of making a composition comprising forming a surface comprising individual sites on a substrate and distributing microspheres on the surface such that the individual sites contain microspheres. The microspheres comprise at least a first and a second subpopulation each comprising a bioactive agent and an identifier binding ligand that will bind a decoder binding ligand such that the identification of the bioactive agent can be elucidated.

In an additional aspect, the invention provides methods of decoding an array composition comprising providing an array composition as outlined above, and adding a plurality of decoding binding ligands to the array composition to identify the location of at least a plurality of the bioactive agents.

In a further aspect, the invention provides methods of determining the presence of a target analyte in a sample. The methods comprise contacting the sample with an array composition as outlined herein, and determining the presence or absence of the target analyte.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is generally based on previous work comprising a bead-based analytic chemistry system in which beads, also termed microspheres, carrying different chemical functionalities are distributed on a substrate comprising a patterned surface of discrete sites that can bind the individual microspheres. Since the beads are generally put onto the substrate randomly, the previous work relied on the incorporation of unique optical signatures, generally fluorescent dyes, that could be used to identify the chemical functionality on any particular bead. This allows the synthesis of the candidate agents (i.e. compounds such as nucleic acids and antibodies) to be divorced from their placement on an array, i.e. the candidate agents may be synthesized on the beads, and then the beads are randomly distributed on a patterned surface. Since the beads are first coded with an optical signature, this means that the array can later be "decoded", i.e. after the array is made, a correlation of the location of an individual site on the array with the bead or candidate agent at that particular site can be made. This means that the beads may be randomly distributed on the array, a fast and inexpensive process as compared to either the in situ synthesis or spotting techniques of the prior art.

However, the drawback to these methods is that for a large array, the system requires a large number of different optical signatures, which may be difficult or time-consuming to utilize. Accordingly, the present invention provides several improvements over these methods, generally directed to methods of coding and decoding the arrays. That is, as will be appreciated by those in the art, the placement of the bioactive agents is generally random, and thus a coding/decoding system is required to identify the

bioactive agent at each location in the array. This may be done in a variety of ways, as is more fully outlined below, and generally includes: a) the use of decoding binding ligands (DBLs), generally directly labeled, that binds to either the bioactive agent or to identifier binding ligands (IBLs) attached to the beads; b) positional decoding, for example by either targeting the placement of beads (for example by using photoactivatible or photocleavable moieties to allow the selective addition of beads to particular locations), or by using either sub-bundles or selective loading of the sites, as are more fully outlined below; c) selective decoding, wherein only those beads that bind to a target are decoded; or d) combinations of any of these. In some cases, as is more fully outlined below, this decoding may occur for all the beads, or only for those that bind a particular target analyte. Similarly, this may occur either prior to or after addition of a target analyte.

Once the identity of the bioactive agent and its location in the array has been fixed, the array is exposed to samples containing the target analytes, although as outlined below, this can be done prior to or during the analysis as well. The target analytes will bind to the bioactive agents as is more fully outlined below, and results in a change in the optical signal of a particular bead.

In the present invention, "decoding" does not rely on the use of optical signatures, but rather on the use of decoding binding ligands that are added during a decoding step. The decoding binding ligands will bind either to a distinct identifier binding ligand partner that is placed on the beads, or to the bioactive agent itself, for example when the beads comprise single-stranded nucleic acids as the bioactive agents. The decoding binding ligands are either directly or indirectly labeled, and thus decoding occurs by detecting the presence of the label. By using pools of decoding binding ligands in a sequential fashion, it is possible to greatly minimize the number of required decoding steps.

Accordingly, the present invention provides array compositions comprising at least a first substrate with a surface comprising individual sites. By "array" herein is meant a plurality of candidate agents in an array format; the size of the array will depend on the composition and end use of the array. Arrays containing from about 2 different bioactive agents (i.e. different beads) to many millions can be made, with very large fiber optic arrays being possible. Generally, the array will comprise from two to as many as a billion or more, depending on the size of the beads and the substrate, as well as the end use of the array, thus very high density, high density, moderate density, low density and very low density arrays may be made. Preferred ranges for very high density arrays are from about 10,000,000 to about 2,000,000,000 (all numbers are per square cm), with from about 100,000,000 to about 1,000,000,000 being preferred. High density arrays range about 100,000 to about 10,000,000, with from about 1,000,000 to about 5,000,000 being particularly preferred. Moderate density arrays range from about 10,000 to about 100,000 being particularly preferred, and from about 20,000 to about 50,000 being especially preferred. Low density arrays are generally less than 10,000, with from about

1,000 to about 5,000 being preferred. Very low density arrays are less than 1,000, with from about 10 to about 1000 being preferred, and from about 100 to about 500 being particularly preferred. In some embodiments, the compositions of the invention may not be in array format; that is, for some embodiments, compositions comprising a single bioactive agent may be made as well. In addition, in some arrays, multiple substrates may be used, either of different or identical compositions. Thus for example, large arrays may comprise a plurality of smaller substrates.

In addition, one advantage of the present compositions is that particularly through the use of fiber optic technology, extremely high density arrays can be made. Thus for example, because beads of 200 μm or less (with beads of 200 nm possible) can be used, and very small fibers are known, it is possible to have as many as 250,000 or more (in some instances, 1 million) different fibers and beads in a 1 mm^2 fiber optic bundle, with densities of greater than 15,000,000 individual beads and fibers (again, in some instances as many as 25-50 million) per 0.5 cm^2 obtainable.

By "substrate" or "solid support" or other grammatical equivalents herein is meant any material that can be modified to contain discrete individual sites appropriate for the attachment or association of beads and is amenable to at least one detection method. As will be appreciated by those in the art, the number of possible substrates is very large. Possible substrates include, but are not limited to, glass and modified or functionalized glass, plastics (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polyurethanes, Teflon, etc.), polysaccharides, nylon or nitrocellulose, resins, silica or silica-based materials including silicon and modified silicon, carbon, metals, inorganic glasses, plastics, optical fiber bundles, and a variety of other polymers. In general, the substrates allow optical detection and do not themselves appreciably fluoresce.

Generally the substrate is flat (planar), although as will be appreciated by those in the art, other configurations of substrates may be used as well; for example, three dimensional configurations can be used, for example by embedding the beads in a porous block of plastic that allows sample access to the beads and using a confocal microscope for detection. Similarly, the beads may be placed on the inside surface of a tube, for flow-through sample analysis to minimize sample volume. Preferred substrates include optical fiber bundles as discussed below, and flat planar substrates such as glass, polystyrene and other plastics and acrylics.

In a preferred embodiment, the substrate is an optical fiber bundle or array, as is generally described in U.S.S.N.s 08/944,850 and 08/519,062, PCT US98/05025, and PCT US98/09163, all of which are expressly incorporated herein by reference. Preferred embodiments utilize preformed unitary fiber optic arrays. By "preformed unitary fiber optic array" herein is meant an array of discrete individual

fiber optic strands that are co-axially disposed and joined along their lengths. The fiber strands are generally individually clad. However, one thing that distinguished a preformed unitary array from other fiber optic formats is that the fibers are not individually physically manipulatable; that is, one strand generally cannot be physically separated at any point along its length from another fiber strand.

5

At least one surface of the substrate is modified to contain discrete, individual sites for later association of microspheres. These sites may comprise physically altered sites, i.e. physical configurations such as wells or small depressions in the substrate that can retain the beads, such that a microsphere can rest in the well, or the use of other forces (magnetic or compressive), or chemically

10   altered or active sites, such as chemically functionalized sites, electrostatically altered sites, hydrophobically/ hydrophilically functionalized sites, spots of adhesive, etc.

The sites may be a pattern, i.e. a regular design or configuration, or randomly distributed. A preferred embodiment utilizes a regular pattern of sites such that the sites may be addressed in the X-Y

15   coordinate plane. "Pattern" in this sense includes a repeating unit cell, preferably one that allows a high density of beads on the substrate. However, it should be noted that these sites may not be discrete sites. That is, it is possible to use a uniform surface of adhesive or chemical functionalities, for example, that allows the association of beads at any position. That is, the surface of the substrate is modified to allow association of the microspheres at individual sites, whether or not those sites are

20   contiguous or non-contiguous with other sites. Thus, the surface of the substrate may be modified such that discrete sites are formed that can only have a single associated bead, or alternatively, the surface of the substrate is modified and beads may go down anywhere, but they end up at discrete sites.

25   In a preferred embodiment, the surface of the substrate is modified to contain wells, i.e. depressions in the surface of the substrate. This may be done as is generally known in the art using a variety of techniques, including, but not limited to, photolithography, stamping techniques, molding techniques and microetching techniques. As will be appreciated by those in the art, the technique used will depend on the composition and shape of the substrate.

30

In a preferred embodiment, physical alterations are made in a surface of the substrate to produce the sites. In a preferred embodiment, the substrate is a fiber optic bundle and the surface of the substrate is a terminal end of the fiber bundle, as is generally described in 08/818,199 and 09/151,877, both of which are hereby expressly incorporated by reference. In this embodiment, wells are made in a

35   terminal or distal end of a fiber optic bundle comprising individual fibers. In this embodiment, the cores of the individual fibers are etched, with respect to the cladding, such that small wells or depressions are formed at one end of the fibers. The required depth of the wells will depend on the size of the

beads to be added to the wells.

Generally in this embodiment, the microspheres are non-covalently associated in the wells, although the wells may additionally be chemically functionalized as is generally described below, cross-linking agents may be used, or a physical barrier may be used, i.e. a film or membrane over the beads.

In a preferred embodiment, the surface of the substrate is modified to contain chemically modified sites, that can be used to associate, either covalently or non-covalently, the microspheres of the invention to the discrete sites or locations on the substrate. "Chemically modified sites" in this context includes, but is not limited to, the addition of a pattern of chemical functional groups including amino groups, carboxy groups, oxo groups and thiol groups, that can be used to covalently attach microspheres, which generally also contain corresponding reactive functional groups; the addition of a pattern of adhesive that can be used to bind the microspheres (either by prior chemical functionalization for the addition of the adhesive or direct addition of the adhesive); the addition of a pattern of charged groups (similar to the chemical functionalities) for the electrostatic association of the microspheres, i.e. when the microspheres comprise charged groups opposite to the sites; the addition of a pattern of chemical functional groups that renders the sites differentially hydrophobic or hydrophilic, such that the addition of similarly hydrophobic or hydrophilic microspheres under suitable experimental conditions will result in association of the microspheres to the sites on the basis of hydroaffinity. For example, the use of hydrophobic sites with hydrophobic beads, in an aqueous system, drives the association of the beads preferentially onto the sites. As outlined above, "pattern" in this sense includes the use of a uniform treatment of the surface to allow association of the beads at discrete sites, as well as treatment of the surface resulting in discrete sites. As will be appreciated by those in the art, this may be accomplished in a variety of ways.

The compositions of the invention further comprise a population of microspheres. By "population" herein is meant a plurality of beads as outlined above for arrays. Within the population are separate subpopulations, which can be a single microsphere or multiple identical microspheres. That is, in some embodiments, as is more fully outlined below, the array may contain only a single bead for each bioactive agent; preferred embodiments utilize a plurality of beads of each type.

By "microspheres" or "beads" or "particles" or grammatical equivalents herein is meant small discrete particles. The composition of the beads will vary, depending on the class of bioactive agent and the method of synthesis. Suitable bead compositions include those used in peptide, nucleic acid and organic moiety synthesis, including, but not limited to, plastics, ceramics, glass, polystyrene, methylstyrene, acrylic polymers, paramagnetic materials, thoria sol, carbon graphite, titanium dioxide, latex or cross-linked dextrans such as Sepharose, cellulose, nylon, cross-linked micelles and Teflon

may all be used. *"Microsphere Detection Guide"* from Bangs Laboratories, Fishers IN is a helpful guide.

5   The beads need not be spherical; irregular particles may be used. In addition, the beads may be porous, thus increasing the surface area of the bead available for either bioactive agent attachment or IBL attachment. The bead sizes range from nanometers, i.e. 100 nm, to millimeters, i.e. 1 mm, with beads from about 0.2 micron to about 200 microns being preferred, and from about 0.5 to about 5 micron being particularly preferred, although in some embodiments smaller beads may be used.

10   It should be noted that a key component of the invention is the use of a substrate/bead pairing that allows the association or attachment of the beads at discrete sites on the surface of the substrate, such that ine beads do not move during the course of the assay.

       Each microsphere comprises a bioactive agent, although as will be appreciated by those in the art,
15   there may be some microspheres which do not contain a bioactive agent, depending on the synthetic methods. By "candidate bioactive agent" or "bioactive agent" or "chemical functionality" or "binding ligand" herein is meant as used herein describes any molecule, *e.g.*, protein, oligopeptide, small organic molecule, coordination complex, polysaccharide, polynucleotide, etc. which can be attached to the microspheres of the invention. It should be understood that the compositions of the invention have
20   two primary uses. In a preferred embodiment, as is more fully outlined below, the compositions are used to detect the presence of a particular target analyte; for example, the presence or absence of a particular nucleotide sequence or a particular protein, such as an enzyme, an antibody or an antigen. In an alternate preferred embodiment, the compositions are used to screen bioactive agents, i.e. drug candidates, for binding to a particular target analyte.

25

       Bioactive agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 100 and less than about 2,500 Daltons. Bioactive agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or
30   carboxyl group, preferably at least two of the functional chemical groups. The bioactive agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Bioactive agents are also found among biomolecules including peptides, nucleic acids, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof. Particularly preferred are nucleic acids and
35   proteins.

       Bioactive agents can be obtained from a wide variety of sources including libraries of synthetic or

natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced

5   libraries and compounds are readily modified through conventional chemical, physical and biochemical means. Known pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification and/or amidification to produce structural analogs.

10   In a preferred embodiment, the bioactive agents are proteins. By "protein" herein is meant at least two covalently attached amino acids, which includes proteins, polypeptides, oligopeptides and peptides. The protein may be made up of naturally occurring amino acids and peptide bonds, or synthetic peptidomimetic structures. Thus "amino acid", or "peptide residue", as used herein means both naturally occurring and synthetic amino acids. For example, homo-phenylalanine, citrulline and

15   norleucine are considered amino acids for the purposes of the invention. The side chains may be in either the (R) or the (S) configuration. In the preferred embodiment, the amino acids are in the (S) or L-configuration. If non-naturally occurring side chains are used, non-amino acid substituents may be used, for example to prevent or retard in vivo degradations.

20   In one preferred embodiment, the bioactive agents are naturally occurring proteins or fragments of naturally occuring proteins. Thus, for example, cellular extracts containing proteins, or random or directed digests of proteinaceous cellular extracts, may be used. In this way libraries of procaryotic and eukaryotic proteins may be made for screening in the systems described herein. Particularly preferred in this embodiment are libraries of bacterial, fungal, viral, and mammalian proteins, with the

25   latter being preferred, and human proteins being especially preferred.

In a preferred embodiment, the bioactive agents are peptides of from about 5 to about 30 amino acids, with from about 5 to about 20 amino acids being preferred, and from about 7 to about 15 being particularly preferred. The peptides may be digests of naturally occurring proteins as is outlined

30   above, random peptides, or "biased" random peptides. By "randomized" or grammatical equivalents herein is meant that each nucleic acid and peptide consists of essentially random nucleotides and amino acids, respectively. Since generally these random peptides (or nucleic acids, discussed below) are chemically synthesized, they may incorporate any nucleotide or amino acid at any position. The synthetic process can be designed to generate randomized proteins or nucleic acids, to allow the

35   formation of all or most of the possible combinations over the length of the sequence, thus forming a library of randomized bioactive proteinaceous agents.

In a preferred embodiment, a library of bioactive agents are used. The library should provide a sufficiently structurally diverse population of bioactive agents to effect a probabilistically sufficient range of binding to target analytes. Accordingly, an interaction library must be large enough so that at least one of its members will have a structure that gives it affinity for the target analyte. Although it is difficult to gauge the required absolute size of an interaction library, nature provides a hint with the immune response: a diversity of 10^7-10^8 different antibodies provides at least one combination with sufficient affinity to interact with most potential antigens faced by an organism. Published in vitro selection techniques have also shown that a library size of 10^7 to 10^8 is sufficient to find structures with affinity for the target. Thus, in a preferred embodiment, at least 10^6, preferably at least 10^7, more preferably at least 10^8 and most preferably at least 10^9 different bioactive agents are simultaneously analyzed in the subject methods. Preferred methods maximize library size and diversity.

In a preferred embodiment, the library is fully randomized, with no sequence preferences or constants at any position. In a preferred embodiment, the library is biased. That is, some positions within the sequence are either held constant, or c.. selected from a limited number of possibilities. For example, in a preferred embodiment, the nucleotides or amino acid residues are randomized within a defined class, for example, of hydrophobic amino acids, hydrophilic residues, sterically biased (either small or large) residues, towards the creation of cysteines, for cross-linking, prolines for SH-3 domains, serines, threonines, tyrosines or histidines for phosphorylation sites, etc., or to purines, etc.

In a preferred embodiment, the bioactive agents are nucleic acids (generally called "probe nucleic acids" or "candidate probes" herein). By "nucleic acid" or "oligonucleotide" or grammatical equivalents herein means at least two nucleotides covalently linked together. A nucleic acid of the present invention will generally contain phosphodiester bonds, although in some cases, as outlined below, nucleic acid analogs are included that may have alternate backbones, comprising, for example, phosphoramide (Beaucage, et al., Tetrahedron, 49(10):1925 (1993) and references therein; Letsinger, J. Org. Chem., 35:3800 (1970); Sprinzl, et al., Eur. J. Biochem., 81:579 (1977); Letsinger, et al., Nucl. Acids Res., 14:3487 (1986); Sawai, et al., Chem. Lett., 805 (1984), Letsinger, et al., J. Am. Chem. Soc., 110:4470 (1988); and Pauwels, et al., Chemica Scripta, 26:141 (1986)), phosphorothioate (Mag, et al., Nucleic Acids Res., 19:1437 (1991); and U.S. Patent No. 5,644,048), phosphorodithioate (Briu, et al., J. Am. Chem. Soc., 111:2321 (1989)), O-methylphophoroamidite linkages (see Eckstein, Oligonucleotides and Analogues: A Practical Approach, Oxford University Press), and peptide nucleic acid backbones and linkages (see Egholm, J. Am. Chem. Soc., 114:1895 (1992); Meier, et al., Chem. Int. Ed. Engl., 31:1008 (1992); Nielsen, Nature, 365:566 (1993); Carlsson, et al., Nature, 380:207 (1996), all of which are incorporated by reference)). Other analog nucleic acids include those with positive backbones (Denpcy, et al., Proc. Natl. Acad. Sci. USA, 92:6097 (1995)); non-ionic backbones (U.S. Patent Nos. 5,386,023; 5,637,684; 5,602,240; 5,216,141; and 4,469,863; Kiedrowshi, et al.,

Angew. Chem. Intl. Ed. English, 30:423 (1991); Letsinger, *et al.*, J. Am. Chem. Soc., 110:4470 (1988);
Letsinger, *et al.*, Nucleosides & Nucleotides, 13:1597 (1994); Chapters 2 and 3, ASC Symposium
Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sanghui and P. Dan Cook;
Mesmaeker, *et al.*, Bioorganic & Medicinal Chem. Lett., 4:395 (1994); Jeffs, *et al.*, J. Biomolecular
5   NMR, 34:17 (1994); Tetrahedron Lett., 37:743 (1996)) and non-ribose backbones, including those
described in U S Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, ASC Symposium
Series 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sanghui and P. Dan Cook.
Nucleic acids containing one or more carbocyclic sugars are also included within the definition of
nucleic acids (see Jenkins, *et al.*, Chem. Soc. Rev., (1995) pp. 169-176). Several nucleic acid analogs
10   are described in Rawls, C & E News, June 2, 1997, page 35. All of these references are hereby
expressly incorporated by reference. These modifications of the ribose-phosphate backbone may be
done to facilitate the addition of additional moieties such as labels, or to increase the stability and half-
life of such molecules in physiological environments; for example, PNA is particularly preferred. In
addition, mixtures of naturally occurring nucleic acids and analogs can be made. Alternatively,
15   mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and
analogs may be made. The nucleic acids may be single stranded or double stranded, as specified, or
contain portions of both double stranded or single stranded sequence. The nucleic acid may be DNA,
both genomic and cDNA, RNA or a hybrid, where the nucleic acid contains any combination of
deoxyribo- and ribo-nucleotides, and any combination of bases, including uracil, adenine, thymine,
20   cytosine, guanine, inosine, xanthanine, hypoxanthanine, isocytosine, isoguanine, and base analogs
such as nitropyrrole and nitroindole, etc.

In a preferred embodiment, the bioactive agents are libraries of clonal nucleic acids, including DNA
and RNA. In this embodiment, individual nucleic acids are prepared, generally using conventional
25   methods (including, but not limited to, propagation in plasmid or phage vectors, amplification
techniques including PCR, etc.). The nucleic acids are preferably arrayed in some format, such as a
microtiter plate format, and beads added for attachment of the libraries.

Attachment of the clonal libraries (or any of the nucleic acids outlined herein) may be done in a variety
30   of ways, as will be appreciated by those in the art, including, but not limited to, chemical or affinity
capture (for example, including the incorporation of derivatized nucleotides such as AminoLink or
biotinylated nucleotides that can then be used to attach the nucleic acid to a surface, as well as affinity
capture by hybridization), cross-linking, and electrostatic attachment, etc.

35   In a preferred embodiment, affinity capture is used to attach the clonal nucleic acids to the beads. For
example, cloned nucleic acids can be derivatized, for example with one member of a binding pair, and
the beads derivatized with the other member of a binding pair. Suitable binding pairs are as described

13

herein for IBL/DBL pairs. For example, the cloned nucleic acids may be biotinylated (for example using enzymatic incorporate of biotinylated nucleotides, for by photoactivated cross-linking of biotin). Biotinylated nucleic acids can then be captured on streptavidin-coated beads, as is known in the art. Similarly, other hapten-receptor combinations can be used, such as digoxigenin and anti-digoxigenin

5   antibodies. Alternatively, chemical groups can he added in the form of derivatized nucleotides, that can them be used to add the nucleic acid to the surface.

Preferred attachments are covalent, although even relatively weak interactions (i.e. non-covalent) can be sufficient to attach a nucleic acid to a surface, if there are multir le sites of attachment per each

10   nucleic acid Thus, for example, electrostatic interactions can be used for attachment, for example by having beads carrying the opposite charge to the bioactive agent.

Similarly, affinity capture utilizing hybridization can be used to attach cloned nucleic acids to beads. For example, as is known in the art, polyA+RNA is routinely captured by hybridization to oligo-dT

15   beads; this may include oligo-dT capture followed by a cross-linking step, such as psoralen crosslinking). If the nucleic acids of interest do not contain a polyA tract, one can be attached by polymerization with terminal transferase, or via ligation of an oligoA linker, as is known in the art.

Alternatively, chemical crosslinking may be done, for example by photoactivated crosslinking of

20   thymidine to reactive groups, as is known in the art.

In general, special methods are required to decode clonal arrays, as is more fully outlined below.

As described above generally for proteins, nucleic acid bioactive agents may be naturally occurring

25   nucleic acids, random nucleic acids, or "biased" random nucleic acids. For example, digests of procaryotic or eukaryotic genomes may be used as is outlined above for proteins.

In general, probes of the present invention are designed to be complementary to a target sequence (either the target analyte sequence of the sample or to other probe sequences, as is described

30   herein), such that hybridization of the target and the probes of the present invention occurs. This complementarily need not be perfect; there may be any number of base pair mismatches that will interfere with hybridization between the target sequence and the single stranded nucleic acids of the present invention. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target

35   sequence. Thus, by "substantially complementary" herein is meant that the probes are sufficiently complementary to the target sequences to hybridize under the selected reaction conditions. High stringency conditions are known in the art; see for example Maniatis et al., *Molecular Cloning: A*

14

Laboratory Manual, 2d Edition, 1989, and *Short Protocols in Molecular Biology*, ed. Ausubel, et al.,
both of which are hereby incorporated by reference. Stringent conditions are sequence-dependent
and will be different in different circumstances. Longer sequences hybridize specifically at higher
temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen, Techniques

5   in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes, "Overview of principles
of hybridization and the strategy of nucleic acid assays" (1993). Generally, stringent conditions are
selected to be about 5-10°C lower than the thermal melting point (T_m) for the specific sequence at a
defined ionic strength pH. The T_m is the temperature (under defined ionic strength, pH and nucleic
acid concentration) at which 50% of the probes complementary to the target hybridize to the target

10   sequence at equilibrium (as the target sequences are present in excess, at T_m, 50% of the probes are
occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than
about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH
7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g. 10 to 50 nucleotides) and
at least about 60°C for long probes (e.g. greater than 50 nucleotides). Stringent conditions may also

15   be achieved with the addition of destabilizing agents such as formamide. In another embodiment, less
stringent hybridization conditions are used; for example, moderate or low stringency conditions may be
used, as are known in the art; see Maniatis and Ausubel, supra, and Tijssen, supra.

The term 'target sequence" or grammatical equivalents herein means a nucleic acid sequence on a

20   single strand of nucleic acid. The target sequence may be a portion of a gene, a regulatory sequence,
genomic DNA, cDNA, RNA including mRNA and rRNA, or others. It may be any length, with the
understanding that longer sequences are more specific. As will be appreciated by those in the art, the
complementary target sequence may take many forms. For example, it may be contained within a
larger nucleic acid sequence, i.e. all or part of a gene or mRNA, a restriction fragment of a plasmid or

25   genomic DNA, among others. As is outlined more fully below, probes are made to hybridize to target
sequences to determine the presence or absence of the target sequence in a sample. Generally
speaking, this term will be understood by those skilled in the art.

In a preferred embodiment, the bioactive agents are organic chemical moieties, a wide variety of which

30   are available in the literature.

In a preferred embodiment, each bead comprises a single type of bioactive agent, although a plurality
of individual bioactive agents are preferably attached to each bead. Similarly, preferred embodiments
utilize more than one microsphere containing a unique bioactive agent; that is, there is redundancy

35   built into the system by the use of subpopulations of microspheres, each microsphere in the
subpopulation containing the same bioactive agent.

15

As will be appreciated by those in the art, the bioactive agents may either be synthesized directly on the beads, or they may be made and then attached after synthesis. In a preferred embodiment, linkers are used to attach the bioactive agents to the beads, to allow both good attachment, sufficient flexibility to allow good interaction with the target molecule, and to avoid undesirable binding reactions.

5

In a preferred embodiment, the bioactive agents are synthesized directly on the beads. As is known in the art, many classes of chemical compounds are currently synthesized on solid supports, including beads, such as peptides, organic moieties, and nucleic acids.

10   In a preferred embodiment, the bioactive agents are synthesized first, and then covalently attached to the beads. As will be appreciated by those in the art, this will be done depending on the composition of the bioactive agents and the beads. The functionalization of solid support surfaces such as certain polymers with chemically reactive groups such as thiols, amines, carboxyls, etc. is generally known in the art. Accordingly, "blank" microspheres may be used that have surface chemistries that facilitate

15   the attachment of the desired functionality by the user. Some examples of these surface chemistries for blank microspheres include, but are not limited to, amino groups including aliphatic and aromatic amines, carboxylic acids, aldehydes, amides, chloromethyl groups, hydrazide, hydroxyl groups, sulfonates and sulfates.

20   These functional groups can be used to add any number of different candidate agents to the beads, generally using known chemistries. For example, candidate agents containing carbohydrates may be attached to an amino-functionalized support; the aldehyde of the carbohydrate is made using standard techniques, and then the aldehyde is reacted with an amino group on the surface. In an alternative embodiment, a sulfhydryl linker may be used. There are a number of sulfhydryl reactive linkers known

25   in the art such as SPDP, maleimides, α-haloacetyls, and pyridyl disulfides (see for example the 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200, incorporated herein by reference) which can be used to attach cysteine containing proteinaceous agents to the support. Alternatively, an amino group on the candidate agent may be used for attachment to an amino group on the surface. For example, a large number of stable bifunctional groups are well

30   known in the art, including homobifunctional and heterobifunctional linkers (see Pierce Catalog and Handbook, pages 155-200). In an additional embodiment, carboxyl groups (either from the surface or from the candidate agent) may be derivatized using well known linkers (see the Pierce catalog). For example, carbodiimides activate carboxyl groups for attack by good nucleophiles such as amines (see Torchilin et al., Critical Rev. Therapeutic Drug Carrier Systems, 7(4):275-308 (1991), expressly

35   incorporated herein). Proteinaceous candidate agents may also be attached using other techniques known in the art, for example for the attachment of antibodies to polymers; see Slinkin et al., Bioconj. Chem. 2:342-348 (1991); Torchilin et al., supra; Trubetskoy et al., Bioconj. Chem. 3:323-327 (1992);

King et al., Cancer Res. 54:6176-6185 (1994); and Wilbur et al., Bioconjugate Chem. 5:220-235 (1994), all of which are hereby expressly incorporated by reference). It should be understood that the candidate agents may be attached in a variety of ways, including those listed above. Preferably, the manner of attachment does not significantly alter the functionality of the candidate agent; that is, the candidate agent should be attached in such a flexible manner as to allow its interaction with a target.

Specific techniques for immobilizing enzymes on microspheres are known in the prior art. In one case, NH_2 surface chemistry microspheres are used. Surface activation is achieved with a 2.5% glutaraldehyde in phosphate buffered saline (10 mM) providing a pH of 6.9. (138 mM NaCl, 2.7 mM, KCl). This is stirred on a stir bed for approximately 2 hours at room temperature. The microspheres are then rinsed with ultrapure water plus 0.01% tween 20 (surfactant) -0.02%, and rinsed again with a pH 7.7 PBS plus 0.01% tween 20. Finally, the enzyme is added to the solution, preferably after being prefiltered using a 0.45µm amicon micropure filter.

In some embodiments, the microspheres may additionally comprise identifier binding ligands for use in certain decoding systems. By "identifier binding ligands" or "IBLs" herein is meant a compound that will specifically bind a corresponding decoder binding ligand (DBL) to facilitate the elucidation of the identity of the bioactive agent attached to the bead. That is, the IBL and the corresponding DBL form a binding partner pair. By "specifically bind" herein is meant that the IBL binds its DBL with specificity sufficient to differentiate between the corresponding DBL and other DBLs (that is, DBLs for other IBLs), or other components or contaminants of the system. The binding should be sufficient to remain bound under the conditions of the decoding step, including wash steps to remove non-specific binding. In some embodiments, for example when the IBLs and corresponding DBLs are proteins or nucleic acids, the dissociation constants of the IBL to its DBL will be less than about 10^{-4}-10^{-6} M^{-1}, with less than about 10^{-5} to 10^{-9} M^{-1} being preferred and less than about 10^{-7} -10^{-9} M^{-1} being particularly preferred.

IBL-DBL binding pairs are known or can be readily found using known techniques. For example, when the IBL is a protein, the DBLs include proteins (particularly including antibodies or fragments thereof (FAbs, etc.)) or small molecules, or vice versa (the IBL is an antibody and the DBL is a protein). Metal ion- metal ion ligands or chelators pairs are also useful. Antigen-antibody pairs, enzymes and substrates or inhibitors, other protein-protein interacting pairs, receptor-ligands, complementary nucleic acids (including nucleic acid molecules that form triple helices), and carbohydrates and their binding partners are also suitable binding pairs. Nucleic acid - nucleic acid binding proteins pairs are also useful, including single-stranded or double-stranded nucleic acid binding proteins, and small molecule nucleic acid binding agents. Similarly, as is generally described in U.S. Patents 5,270,163, 5,475,096, 5,567,588, 5,595,877, 5,637,459, 5,683,867,5,705,337, and related patents, hereby

17

incorporated by reference, nucleic acid "aptamers" can be developed for binding to virtually any target; such an aptamer-target pair can be used as the IBL-DBL pair. Similarly, there is a wide body of literature relating to the development of binding pairs based on combinatorial chemistry methods.

5   In a preferred embodiment, the IBL is a molecule whose color or luminescence properties change in the presence of a selectively-binding DBL.

       In one embodiment, the DBL may be attached to a bead, i.e. a "decoder bead", that may carry a label such as a fluorophore.

10

       In a preferred embodiment, the IBL-DBL pair comprise substantially complementary single-stranded nucleic acids. In this embodiment, the binding ligands can be referred to as "identifier probes" and "decoder probes". Generally, the identifier and decoder probes range from about 4 basepairs in length to about 1000, with from about 6 to about 100 being preferred, and from about 8 to about 40 being

15   particularly preferred. What is important is that the probes are long enough to be specific, i.e. to distinguish between different IBL-DBL pairs, yet short enough to allow both a) dissociation, if necessary, under suitable experimental conditions, and b) efficient hybridization.

       In a preferred embodiment, as is more fully outlined below, the IBLs do not bind to DBLs. Rather, the

20   IBLs are used as identifier moieties ("IMs") that are identified directly, for example through the use of mass spectroscopy.

       Alternatively, in a preferred embodiment, the IBL and the bioactive agent are the same moiety; thus, for example, as outlined herein, particularly when no optical signatures are used, the bioactive agent

25   can serve as both the identifier and the agent. For example, in the case of nucleic acids, the bead-bound probe (which serves as the bioactive agent) can also bind decoder probes, to identify the sequence of the probe on the bead. Thus, in this embodiment, the DBLs bind to the bioactive agents. This is particularly useful as this embodiment can give information about the array or the assay in addition to decoding. For example, as is more fully described below, the use of the DBLs allows array

30   calibration and assay development. This may be done even if the DBLs are not used as such; for example in non-random arrays, the use of these probe sets can allow array calibration and assay development even if decoding is not required.

       In a preferred embodiment, the microspheres do not contain an optical signature. That is, as outlined

35   in U.S.S.N.s 08/818,199 and 09/151,877, previous work had each subpopulation of microspheres comprising a unique optical signature or optical tag that is used to identify the unique bioactive agent of that subpopulation of microspheres; that is, decoding utilizes optical properties of the beads such

that a bead comprising the unique optical signature may be distinguished from beads at other locations with different optical signatures. Thus the previous work assigned each bioactive agent a unique optical signature such that any microspheres comprising that bioactive agent are identifiable on the basis of the signature. These optical signatures comprised dyes, usually chromophores or

5   fluorophores, that were entrapped or attached to the beads themselves. Diversity of optical signatures utilized different fluorochromes, different ratios of mixtures of fluorochromes, and different concentrations (intensities) of fluorochromes.

Thus, the present invention does not rely solely on the use of optical properties to decode the arrays.
10   However, as will be appreciated by those in the art, it is possible in some embodiments to utilize optical signatures as an additional coding method, in conjunction with the present system. Thus, for example, as is more fully outlined below, the size of the array may be effectively increased while using a single set of decoding moieties in several ways, one of which is the use in combination with optical signatures one beads. Thus, for example, using one "set" of decoding molecules, the use of two

15   populations of beads, one with an optical signature and one without, allows the effective doubling of the array size. The use of multiple optical signatures similarly increases the possible size of the array.

In a preferred embodiment, each subpopulation of beads comprises a plurality of different IBLs. By using a plurality of different IBLs to encode each bioactive agent, the number of possible unique codes is substantially increased. That is, by using one unique IBL per bioactive agent, the size of the array
20   will be the number of unique IBLs (assuming no "reuse" occurs, as outlined below). However, by using a plurality of different IBLs per bead, n, the size of the array can be increased to 2^n, when the presence or absence of each IBL is used as the indicator. For example, the assignment of 10 IBLs per bead generates a 10 bit binary code, where each bit can be designated as "1" (IBL is present) or
25   "0" (IBL is absent). A 10 bit binary code has 2^{10} possible variants However, as is more fully discussed below, the size of the array may be further increased if another parameter is included such as concentration or intensity; thus for example, if two different concentrations of the IBL are used, then the array size increases as 3^n. Thus, in this embodiment, each individual bioactive agent in the array is assigned a combination of IBLs, which can be added to the beads prior to the addition of the bioactive
30   agent, after, or during the synthesis of the bioactive agent, i.e. simultaneous addition of IBLs and bioactive agent components.

Alternatively, when the bioactive agent is a polymer of different residues, i.e. when the bioactive agent is a protein or nucleic acid, the combination of different IBLs can be used to elucidate the sequence of
35   the protein or nucleic acid.

Thus, for example, using two different IBLs (IBL1 and IBL2), the first position of a nucleic acid can be

elucidated: for example, adenosine can be represented by the presence of both IBL1 and IBL2; thymidine can be represented by the presence of IBL1 but not IBL2, cytosine can be represented by the presence of IBL2 but not IBL1, and guanosine can be represented by the absence of both. The second position of the nucleic acid can be done in a similar manner using IBL3 and IBL4; thus, the

5   presence of IBL1, IBL2, IBL3 and IBL4 gives a sequence of AA; IBL1, IBL2, and IBL3 shows the sequence AT; IBL1, IBL3 and IBL4 gives the sequence TA, etc. The third position utilizes IBL5 and IBL6. etc. In this way, the use of 20 different identifiers can yield a unique code for every possible 10-mer.

10   The system is similar for proteins but requires a larger number of different IBLs to identify each position, depending on the allowed diversity at each position. Thus for example, if every amino acid is allowed at every position, five different IBLs are required for each position. However, as outlined above, for example when using random peptides as the bioactive agents, there may be bias built into the system; not all amino acids may be present at all positions, and some positions may be preset;

15   accordingly, it may be possible to utilize four different IBLs for each amino acid.

In this way, a sort of "bar code" for each sequence can be constructed; the presence or absence of each distinct IBL will allow the identification of each bioactive agent.

20   In addition, the use of different concentrations or densities of IBLs allows a "reuse" of sorts. If, for example, the bead comprising a first agent has a 1X concentration of IBL, and a second bead comprising a second agent has a 10X concentration of IBL, using saturating concentrations of the corresponding labelled DBL allows the user to distinguish between the two beads.

25   Once the microspheres comprising the candidate agents and the unique IBLs are generated, they are added to the substrate to form an array. It should be noted that while most of the methods described herein add the beads to the substrate prior to the assay, the order of making, using and decoding the array can vary. For example, the array can be made, decoded, and then the assay done. Alternatively, the array can be made, used in an assay, and then decoded; this may find particular use

30   when only a few beads need be decoded. Alternatively, the beads can be added to the assay mixture, i.e. the sample containing the target analytes, prior to the addition of the beads to the substrate; after addition and assay, the array may be decoded. This is particularly preferred when the sample comprising the beads is agitated or mixed; this can increase the amount of target analyte bound to the beads per unit time, and thus (in the case of nucleic acid assays) increase the hybridization kinetics.

35   This may find particular use in cases where the concentration of target analyte in the sample is low; generally, for low concentrations, long binding times must be used.

In addition, adding the beads to the assay mixture can allow sorting or selection. For example, a large library of beads may be added to a sample, and only those beads that bind the sample may be added to the substrate. For example, if the target analyte is fluorescently labeled (either directly (for example by the incorporation of labels into nucleic acid amplification reactions) or indirectly (for example via the use of sandwich assays)), beads that exhibit fluorescence as a result of target analyte binding can be sorted via Fluorescence Activated Cell Sorting (FACS) and only these beads added to an array and subsequently decoded. Similarly, the sorting may be accomplished through affinity techniques; affinity columns comprising the target analytes can be made, and only those beads which bind are used on the array. Similarly, two bead systems can be used; for example, magnetic beads comprising the target analytes can be used to "pull out" those beads that will bind to the targets, followed by subsequent release of the magnetic beads (for example via temperature elevation) and addition to an array.

In general, the methods of making the arrays and of decoding the arrays is done to maximize the number of different candidate agents that can be uniquely encoded. The compositions of the invention may be made in a variety of ways. In general, the arrays are made by adding a solution or slurry comprising the beads to a surface containing the sites for association of the beads. This may be done in a variety of buffers, including aqueous and organic solvents, and mixtures. The solvent can evaporate, and excess beads removed.

In a preferred embodiment, when non-covalent methods are used to associate the beads to the array, a novel method of loading the beads onto the array is used. This method comprises exposing the array to a solution of particles (including microspheres and cells) and then applying energy, e.g. agitating or vibrating the mixture. This results in an array comprising more tightly associated particles, as the agitation is done with sufficient energy to cause weakly-associated beads to fall off (or out, in the case of wells). These sites are then available to bind a different bead. In this way, beads that exhibit a high affinity for the sites are selected. Arrays made in this way have two main advantages as compared to a more static loading: first of all, a higher percentage of the sites can be filled easily, and secondly, the arrays thus loaded show a substantial decrease in bead loss during assays. Thus, in a preferred embodiment, these methods are used to generate arrays that have at least about 50% of the sites filled, with at least about 75% being preferred, and at least about 90% being particularly preferred. Similarly, arrays generated in this manner preferably lose less than about 20% of the beads during an assay, with less than about 10% being preferred and less than about 5% being particularly preferred.

In this embodiment, the substrate comprising the surface with the discrete sites is immersed into a solution comprising the particles (beads, cells, etc.). The surface may comprise wells, as is described

21

herein, or other types of sites on a patterned surface such that there is a differential affinity for the sites. This diffrenetial affinity results in a competitive process, such that particles that will associate more tightly are selected. Preferably. the entire surface to be "loaded" with beads is in fluid contact with the solution. This solution is generally a slurry ranging from about 10,000:1 beads:solution

5   (vol:vol) to 1:1. Generally, the solution can comprise any number of reagents, including aqueous buffers, organic solvents, salts, other reagent components, etc. In addition, the solution preferably comprises an excess of beads; that is, there are more beads than sites on the array. Preferred embodiments utilize two-fold to billion-fold excess of beads.

10   The immersion can mimic the assay conditions; for example, if the array is to be "dipped" from above into a microtiter plate comprising samples, this configuration can be repeated for the loading, thus minimizing the beads that are likely to fall out due to gravity.

       Once the surface has been immersed, the substrate, the solution, or both are subjected to a
15   competitive process, whereby the particles with lower affinity can be disassociated from the substrate and replaced by particles exhibiting a higher affinity to the site. This competitive process is done by the introduction of energy, in the form of heat, sonication, stirring or mixing, vibrating or agitating the solution or substrate, or both.

20   A preferred embodiment utilizes agitation or vibration. In general, the amount of manipulation of the substrate is minimized to prevent damage to the array; thus, preferred embodiments utilize the agitation of the solution rather than the array, although either will work. As will be appreciated by those in the art, this agitation can take on any number of forms, with a preferred embodiment utilizing microtiter plates comprising bead solutions being agitated using microtiter plate shakers.
25
       The agitation proceeds for a period of time sufficient to load the array to a desired fill. Depending on the size and concentration of the beads and the size of the array, this time may range from about 1 second to days, with from about 1 minute to about 24 hours being preferred.

30   It should be noted that not all sites of an array may comprise a bead; that is, there may be some sites on the substrate surface which are empty. In addition, there may be some sites that contain more than one bead, although this is not preferred.

       In some embodiments, for example when chemical attachment is done, it is possible to associate the
35   beads in a non-random or ordered way. For example, using photoactivatible attachment linkers or photoactivatible adhesives or masks, selected sites on the array may be sequentially rendered suitable for attachment, such that defined populations of beads are laid down.

22

The arrays of the present invention are constructed such that information about the identity of the candidate agent is built into the array, such that the random deposition of the beads in the fiber wells can be "decoded" to allow identification of the candidate agent at all positions. This may be done in a variety of ways, and either before, during or after the use of the array to detect target molecules.

Thus, after the array is made, it is "decoded" in order to identify the location of one or more of the bioactive agents, i.e. each subpopulation of beads, on the substrate surface.

In a preferred embodiment, a selective decoding system is used. In this case, only those microspheres exhibiting a change in the optical signal as a result of the binding of a target analyte are decoded. This is commonly done when the number of "hits", i.e. the number of sites to decode, is generally low. That is, the array is first scanned under experimental conditions in the absence of the target analytes. The sample containing the target analytes is added, and only those locations exhibiting a change in the optical signal are decoded. For example, the beads at either the positive or negative signal locations may be either selectively tagged or released from the array (for example through the use of photocleavable linkers), and subsequently sorted or enriched in a fluorescence-activated cell sorter (FACS). That is, either all the negative beads are released, and then the positive beads are either released or analyzed in situ, or alternatively all the positives are released and analyzed. Alternatively, the labels may comprise halogenated aromatic compounds, and detection of the label is done using for example gas chromatography, chemical tags, isotopic tags, or mass spectral tags.

As will be appreciated by those in the art, this may also be done in systems where the array is not decoded; i.e. there need not ever be a correlation of bead composition with location. In this embodiment, the beads are loaded on the array, and the assay is run. The "positives", i.e. those beads displaying a change in the optical signal as is more fully outlined below, are then "marked" to distinguish or separate them from the "negative" beads. This can be done in several ways, preferably using fiber optic arrays. In a preferred embodiment, each bead contains a fluorescent dye. After the assay and the identification of the "positives" or "active beads", light is shown down either only the positive fibers or only the negative fibers, generally in the presence of a light-activated reagent (typically dissolved oxygen). In the former case, all the active beads are photobleached. Thus, upon non-selective release of all the beads with subsequent sorting, for example using a fluorescence activated cell sorter (FACS) machine, the non-fluorescent active beads can be sorted from the fluorescent negative beads. Alternatively, when light is shown down the negative fibers, all the negatives are non-fluorescent and the the postives are fluorescent, and sorting can proceed. The characterization of the attached bioactive agent may be done directly, for example using mass

23

spectroscopy.

Alternatively, the identification may occur through the use of identifier moieties ("IMs"), which are
similar to IBLs but need not necessarily bind to DBLs. That is, rather than elucidate the structure of

5   the bioactive agent directly, the composition of the IMs may serve as the identifier. Thus, for example,
a specific combination of IMs can serve to code the bead, and be used to identify the agent on the
bead upon release from the bead followed by subsequent analysis, for example using a gas
chromatograph or mass spectroscope.

10   Alternatively, rather than having each bead contain a fluorescent dye, each bead comprises a non-
fluorescent precursor to a fluorescent dye. For example, using photocleavable protecting groups,
such as certain ortho-nitrobenzyl groups, on a fluorescent molecule, photoactivation of the
fluorochrome can be done. After the assay, light is shown down again either the "positive" or the
"negative" fibers, to distinguish these populations. The illuminated precursors are then chemically

15   converted to a fluorescent dye. All the beads are then released from the array, with sorting, to form
populations of fluorescent and non-fluorescent beads (either the positives and the negatives or vice
versa).

In an alternate preferred embodiment, the sites of association of the beads (for example the wells)

20   include a photopolymerizable reagent, or the photopolymerizable agent is added to the assembled
array. After the test assay is run, light is shown down again either the "positive" or the "negative"
fibers, to distinquish these populations. As a result of the irradiation, either all the positives or all the
negatives are polymerized and trapped or bound to the sites, while the other population of beads can
be released from the array.

25

In a preferred embodiment, the location of every bioactive agent is determined using decoder binding
ligands (DBLs). As outlined above, DBLs are binding ligands that will either bind to identifier binding
ligands, if present, or to the bioactive agents themselves, preferably when the bioactive agent is a
nucleic acid or protein.

30

In a preferred embodiment, as outlined above, the DBL binds to the IBL.

In a preferred embodiment, the bioactive agents are single-stranded nucleic acids and the DBL is a
substantially complementary single-stranded nucleic acid that binds (hybridizes) to the bioactive agent,

35   termed a decoder probe herein. A decoder probe that is substantially complementary to each
candidate probe is made and used to decode the array. In this embodiment, the candidate probes and
the decoder probes should be of sufficient length (and the decoding step run under suitable

conditions) to allow specificity; i.e. each candidate probe binds to its corresponding decoder probe with sufficient specificity to allow the distinction of each candidate probe.

5

In a preferred embodiment, the DBLs are either directly or indirectly labeled. By "labeled" herein is meant that a compound has at least one element, isotope or chemical compound attached to enable the detection of the compound. In general, labels fall into three classes: a) isotopic labels, which may be radioactive or heavy isotopes; b) magnetic, electrical, thermal; and c) colored or luminescent dyes; although labels include enzymes and particles such as magnetic particles as well. Preferred labels include luminescent labels. In a preferred embodiment, the DBL is directly labeled, that is, the DBL

10   comprises a label. In an alternate embodiment, the DBL is indirectly labeled; that is, a labeling binding ligand (LBL) that will bind to the DBL is used. In this embodiment, the labeling binding ligand-DBL pair can be as described above for IBL-DBL pairs. Suit: . e labels include, but are not limited to, fluorescent lanthanide complexes, including those of Europium and Terbium, fluorescein, rhodamine, tetramethylrhodamine, eosin, erythrosin, coumarin, methyl-coumarins, pyrene, Malacite green,

15   stilbene, Lucifer Yellow, Cascade Blue™, Texas Red, FITC, PE, cy3, cy5 and others described in the 6th Edition of the Molecular Probes Handbook by Richard P. Haugland, hereby expressly incorporated by reference.

In one embodiment, the label is a molecule whose color or luminescence properties change in the

20   presence of the IBL, due to a change in the local environment. For example, the label may be: (1) a fluorescent pH indicator whose emission intensity changes with pH; (2) a fluorescent ion indicator, whose emission properties change with ion concentration; or (3) a fluorescent molecule such as an ethidium salt whose fluorescence intensity increases in hydrophobic environments.

25   Accordingly, the identification of the location of the individual beads (or subpopulations of beads) is done using one or more decoding steps comprising a binding between the labeled DBL and either the IBL or the bioactive agent (i.e. a hybridization between the candidate probe and the decoder probe when the bioactive agent is a nucleic acid). After decoding, the DBLs can be removed and the array can be used; however, in some circumstances, for example when the DBL binds to an IBL and not to

30   the bioactive agent, the removal of the DBL is not required (although it may be desirable in some circumstances). In addition, as outlined herein, decoding may be done either before the array is used in an assay, dur..y the assay, or after the assay.

In one embodiment, a single decoding step is done. In this embodiment, each DBL is labeled with a

35   unique label, such that the the number of unique labels is equal to or greater than the number of bioactive agents (although in some cases, "reuse" of the unique labels can be done, as described herein; similarly, minor variants of candidate probes can share the same decoder, if the variants are

25

encoded in another dimension, i.e. in the bead size or label). For each bioactive agent or IBL, a DBL is made that will specifically bind to it and contains a unique label, for example one or more fluorochromes. Thus, the identity of each DBL, both its composition (i.e. its sequence when it is a nucleic acid) and its label, is known. Then, by adding the DBLs to the array containing the bioactive

5   agents under conditions which allow the formation of complexes (termed hybridization complexes when the components are nucleic acids) between the DBLs and either the bioactive agents or the IBLs, the location of each DBL can be elucidated. This allows the identification of the location of each bioactive agent; the random array has been decoded. The DBLs can then be removed, if necessary, and the target sample applied.

10

In a preferred embodiment, the number of unique labels is less than the number of unique bioactive agents, and thus a sequential series of decoding steps are used. To facilitate the discussion, this embodiment is explained for nucleic acids, although other types of bioactive agents and DBLs are useful as well. In this embodiment, decoder probes are divided into n sets for decoding. The number

15   of sets corresponds to the number of unique tags. Each decoder probe is labeled in n separate reactions with n distinct tags. All the decoder probes share the same n tags. Each pool of decoders contains only one of the n tag versions of each decoder, and no two decoder probes have the same sequence of tags across all the pools. The number of pools required for this to be true is determined by the number of decoder probes and the n. Hybridization of each pool to the array generates a signal

20   at every address comprising an IBL. The sequential hybridization of each pool in turn will generate a unique, sequence-specific code for each candidate probe. This identifies the candidate probe at each address in the array. For example, if four tags are used, then 4 X n sequential hybridizations can ideally distinguish 4^n sequences, although in some cases more steps may be required. After the hybridization of each pool, the hybrids are denatured and the decoder probes removed, so that the

25   probes are rendered single-stranded for the next hybridization (although it is also possible to hybridize limiting amounts of target so that the available probe is not saturated. Sequential hybridizations can be carried out and analyzed by subtracting pre-existing signal from the previous hybridization).

An example is illustrative. Assuming an array of 16 probe nucleic acids (numbers 1-16), and four

30   unique tags (four different fluors, for example; labels A-D). Decoder probes 1-16 are made that correspond to the probes on the beads. The first step is to label decoder probes 1-4 with tag A, decoder probes 5-8 with tag B, decoder probes 9-12 with tag C, and decoder probes 13-16 with tag D. The probes are mixed and the pool is contacted with the array containing the beads with the attached candidate probes. The location of each tag (and thus each decoder and candidate probe pair) is then

35   determined. The first set of decoder probes are then removed. A second set is added, but this time, decoder probes 1, 5, 9 and 13 are labeled with tag A, decoder probes 2, 6, 10 and 14 are labeled with tag B, decoder probes 3, 7, 11 and 15 are labeled with tag C, and decoder probes 4, 8, 12 and 16 are

labeled with tag D. Thus, those beads that contained tag A in both decoding steps contain candidate probe 1; tag A in the first decoding step and tag B in the second decoding step contain candidate probe 2; tag A in the first decoding step and tag C in the second step contain candidate probe 3; etc. As will be appreciated by those in the art, the decoder probes can be made in any order and added in

5   any order.

In one embodiment, the decoder probes are labeled in situ; that is, they need not be labeled prior to the decoding reaction. In this embodiment, the incoming decoder probe is shorter than the candidate probe, creating a 5' "overhang" on the decoding probe. The addition of labeled ddNTPs (each labeled

10   with a unique tag) and a polymerase will allow the addition of the tags in a sequence specific manner, thus creating a sequence-specific pattern of signals. Similarly, other modifications can be done, including ligation, etc.

In addition, since the size of the array will be set by the number of unique decoding binding ligands, it

15   is possible to "reuse" a set of unique DBLs to allow for a greater number of test sites. This may be done in several ways; for example, by using some subpopulations that comprise optical signatures. Similarly, the use of a positional coding scheme within an array; different sub-bundles may reuse the set of DBLs. Similarly, one embodiment utilizes bead size as a coding modality, thus allowing the reuse of the set of unique DBLs for each bead size. Alternatively, sequential partial loading of arrays

20   with beads can also allow the reuse of DBLs. Furthermore, "code sharing" can occur as well.

In a preferred embodiment, the DBLs may be reused by having some subpopulations of beads comprise optical signatures. In a preferred embodiment, the optical signature is generally a mixture of reporter dyes, preferably fluoroscent. By varying both the composition of the mixture (i.e. the ratio of

25   one dye to another) and the concentration of the dye (leading to differences in signal intensity), matrices of unique optical signatures may be generated. This may be done by covalently attaching the dyes to the surface of the beads, or alternatively, by entrapping the dye within the bead. The dyes may be chromophores or phosphors but are preferably fluorescent dyes, which due to their strong signals provide a good signal-to-noise ratio for decoding. Suitable dyes for use in the invention include

30   those listed for labeling DBLs, above.

In a preferred embodiment, the encoding can be accomplished in a ratio of at least two dyes, although more encoding dimensions may be added in the size of the beads, for example. In addition, the labels are distinguishable from one another; thus two different labels may comprise different molecules (i.e.

35   two different fluors) or, alternatively, one label at two different concentrations or intensity.

In a preferred embodiment, the dyes are covalently attached to the surface of the beads. This may be

27

done as is generally outlined for the attachment of the bioactive agents, using functional groups on the surface of the beads. As will be appreciated by those in the art, these attachments are done to minimize the effect on the dye.

5   In a preferred embodiment, the dyes are non-covalently associated with the beads, generally by entrapping the dyes in the pores of the beads.

Additionally, encoding in the ratios of the two or more dyes, rather than single dye concentrations, is preferred since it provides insensitivity to the intensity of light used to interrogate the reporter dye's
10   signature and detector sensitivity.

In a preferred embodiment, a spatial or positional coding system is done. In this embodiment, there are sub-bundles or subarrays (i.e. portions of the total array) that are utilized. By analogy with the telephone system, each subarray is an "area code", that can have the same labels (i.e. telephone
15   numbers) of other subarrays, that are separated by virtue of the location of the subarray. Thus, for example, the same unique labels can be reused from bundle to bundle. Thus, the use of 50 unique labels in combination with 100 different subarrays can form an array of 5000 different bioactive agents. In this embodiment, it becomes important to be able to identify one bundle from another; in general, this is done either manually or through the use of marker beads; these can be beads containing
20   unique tags for each subarray, or the use of the same marker bead in differing amounts, or the use of two or more marker beads in different ratios.

In alternative embodiments, additional encoding parameters can be added, such as microsphere size. For example, the use of different size beads may also allow the reuse of sets of DBLs; that is, it is
25   possible to use microspheres of different sizes to expand the encoding dimensions of the microspheres. Optical fiber arrays can be fabricated containing pixels with different fiber diameters or cross-sections; alternatively, two or more fiber optic bundles, each with different cross-sections of the individual fibers, can be added together to form a larger bundle; or, fiber optic bundles with fiber of the same size cross-sections can be used, but just with different sized beads. With different diameters, the
30   largest wells can be filled with the largest microspheres and then moving onto progressively smaller microspheres in the smaller wells until all size wells are then filled. In this manner, the same dye ratio could be used to encode microspheres of different sizes thereby expanding the number of different oligonucleotide sequences or chemical functionalities present in the array. Although outlined for fiber optic substrates, this as well as the other methods outlined herein can be used with other substrates
35   and with other attachment modalities as well.

In a preferred embodiment, the coding and decoding is accomplished by sequential loading of the

28

microspheres into the array. As outlined above for spatial coding, in this embodiment, the optical signatures can be "reused". In this embodiment, the library of microspheres each comprising a different bioactive agent (or the subpopulations each comprise a different bioactive agent), is divided into a plurality of sublibraries; for example, depending on the size of the desired array and the number

5   of unique tags, 10 sublibraries each comprising roughly 10% of the total library may be made, with each sublibrary comprising roughly the same unique tags. Then, the first sublibrary is added to the fiber optic bundle comprising the wells, and the location of each bioactive agent is determined, generally through the use of DBLs. The second sublibrary is then added, and the location of each bioactive agent is again determined. The signal in this case will comprise the signal from the "first"

10   DBL and the "second" DBL; by comparing the two matrices the location of each bead in each sublibrary can be determined. Similarly, adding the third, fourth, etc. sublibraries sequentially will allow the array to be filled.

In a preferred embodiment, codes can be "shared" in several ways. In a first embodiment, a single

15   code (i.e. IBL/DBL pair) can be assigned to two or more agents if the target analytes different sufficiently in their binding strengths. For example, two nucleic acid probes used in an mRNA quantitation assay can share the same code if the ranges of their hybridization signal intensities do not overlap. This can occur, for example, when one of the target sequences is always present at a much higher concentration than the other. Alternatively, the two target sequences might always be present

20   at a similar concentration, but differ in hybridization efficiency.

Alternatively, a single code can be assigned to multiple agents if the agents are functionally equivalent. For example, if a set of oligonucleotide probes are designed with the common purpose of detecting the presence of a particular gene, then the probes are functionally equivalent, even though they may differ

25   in sequence. Similarly, if classes or "families" of analytes are desired, all probes for different members of a class such as kinases or G-protein coupled receptors could share a code. Similarly, an array of this type could be used to detect homologs of known genes. In this embodiment, each gene is represented by a heterologous set of probes, hybridizing to different regions of the gene (and therefore differing in sequence). The set of probes share a common code. If a homolog is present, it might

30   hybridize to some but not all of the probes. The level of homology might be indicated by the fraction of probes hybridizing, as well as the average hybridization intensity. Similarly, multiple antibodies to the same protein could all share the same code.

In a preferred embodiment, decoding of self-assembled random arrays is done on the bases of pH

35   titration. In this embodiment, in addition to bioactive agents, the beads comprise optical signatures, wherein the optical signatures are generated by the use of pH-responsive dyes (sometimes referred to herein as "pH dyes") such as fluorophores. This embodiment is similar to that outlined in PCT

US98/05025 and U.S.S.N. 09/151,877, both of which are expressly incorporated by reference, except that the dyes used in the present invention exhibits changes in fluorescence intensity (or other properties) when the solution pH is adjusted from below the pKa to above the pKa (or vice versa). In a preferred embodiment, a set of pH dyes is used, each with a different pKa, preferably separated by at least 0.5 pH units. Preferred embodiments utilize a pH dye set of pKa's of 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5, 10.0, 10.5, 11, and 11.5. Each bead can contain any subset of the pH dyes, and in this way a unique code for the bioactive agent is generated. Thus, the decoding of an array is achieved by titrating the array from pH 1 to pH 13, and measuring the fluorescence signal from each bead as a function of solution pH

Thus, the present invention provides array compositions comprising a substrate with a surface comprising discrete sites. A population of microspheres is distributed on the sites, and the population comprises at least a first and a second subpopulation. Each subpopulation comprises a bioactive agent, and, in addition, at least one optical dye with a given pKa. The pKas of the different optical dyes are different.

In a preferred embodiment, when for example the array comprises cloned nucleic acids, there are several methods that can be used to decode the arrays. In a preferred embodiment, when some sequence information about the cloned nucleic acids is known, specific decoding probes can be made as is generally outlined herein.

In a preferred embodiment, "random" decoding probes can be made. By sequential hybridizations or the use of multiple labels, as is outlined above, a unique hybridization pattern can be generated for each sensor element. This allows all the beads representing a given clone to be identified as belonging to the same group. In general, this is done by using random or partially degenerate decoding probes, that bind in a sequence-dependent but not highly sequence-specific manner. The process can be repeated a number of times, each time using a different labeling entity, to generate a different pattern of signals based on quasi-specific interactions. In this way, a unique optical signature is eventually built up for each sensor element. By applying pattern recognition or clustering algorithms to the optical signatures, the beads can be grouped into sets that share the same signature (i.e. carry the same probes).

In order to identify the actual sequence of the clone itself, additional procedures are required; for example, direct sequencing can be done. By using an ordered array containing the clones, such as a spotted cDNA array, a "key" can be generated that links a hybridization pattern to a specific clone whose position in the set is known. In this way the clone can be recovered and further characterized.

30

Alternatively, clone arrays can be decoded using binary decoding with vector tags. For example, partially randomized oligos are cloned into a nucleic acid vector (e.g. plasmid, phage, etc.). Each oligonucleotide sequence consists of a subset of a limited set of sequences. For example, if the limites set comprises 10 sequences, each oligonucleotide may have some subset (or all of the 10)

5   sequences. Thus each of the 10 sequences can be present or absent in the oligonucleotide. Therefore, there are 2^{10} or 1,024 possible combinations. The sequences may overlap, and minor variants can also be represented (e.g. A, C, T and G substitutions) to increase the number of possible combinations. A nucleic acid library is cloned into a vector containing the random code sequences. Alternatively, other methods such as PCR can be used to add the tags. In this way it is possible to

10   use a small number of oligo decoding probes to decode an array of clones.

Once made, the compositions of the invention find use in a number of applications. In a preferred embodiment, the compositions are used to probe a sample solution for the presence or absence of a target analyte; including the quantification of the amount of target analyte present. By "target analyte"

15   or "analyte" or grammatical equivalents herein is meant any atom, molecule, ion, molecular ion, compound or particle to be either detected or evaluated for binding partners. As will be appreciated by those in the art, a large number of analytes may be used in the present invention; basically, any target analyte can be used which binds a bioactive agent or for which a binding partner (i.e. drug candidate) is sought.

20

Suitable analytes include organic and inorganic molecules, including biomolecules. When detection of a target analyte is done, suitable target analytes include, but are not limited to, an environmental pollutant (including pesticides, insecticides, toxins, etc.); a chemical (including solvents, polymers, organic materials, etc.); therapeutic molecules (including therapeutic and abused drugs, antibiotics,

25   etc.); biomolecules (including hormones, cytokines, proteins, nucleic acids, lipids, carbohydrates, cellular membrane antigens and receptors (neural, hormonal, nutrient, and cell surface receptors) or their ligands, etc); whole cells (including procaryotic (such as pathogenic bacteria) and eukaryotic cells, including mammalian tumor cells); viruses (including retroviruses, herpesviruses, adenoviruses, lentiviruses, etc.); and spores; etc. Particularly preferred analytes are nucleic acids and proteins.

30

In a preferred embodiment, the target analyte is a protein. As will be appreciated by those in the art, there are a large number of possible proteinaceous target analytes that may be detected or evaluated for binding partners using the present invention. Suitable protein target analytes include, but are not limited to, (1) immunoglobulins; (2) enzymes (and other proteins); (3) hormones and cytokines (many

35   of which serve as ligands for cellular receptors); and (4) other proteins.

In a preferred embodiment, the target analyte is a nucleic acid. These assays find use in a wide

variety of applications.

In a preferred embodiment, the probes are used in genetic diagnosis. For example, probes can be made using the techniques disclosed herein to detect target sequences such as the gene for
5   nonpolyposis colon cancer, the BRCA1 breast cancer gene, P53, which is a gene associated with a variety of cancers, the Apo E4 gene that indicates a greater risk of Alzheimer's disease, allowing for easy presymptomatic screening of patients, mutations in the cystic fibrosis gene, cytochrome p450s or any of the others well known in the art.

10   In an additional embodiment, viral and bacterial detection is done using the complexes of the invention. In this embodiment, probes are designed to detect target sequences from a variety of bacteria and viruses. For example, current blood-screening techniques rely on the detection of anti-HIV antibodies. The methods disclosed herein allow for direct screening of clinical samples to detect HIV nucleic acid sequences, particularly highly conserved HIV sequences. In addition, this allows
15   direct monitoring of circulating virus within a patient as an improved method of assessing the efficacy of anti-viral therapies. Similarly, viruses associated with leukemia, HTLV-I and HTLV-II, may be detected in this way. Bacterial infections such as tuberculosis, chlamydia and other sexually transmitted diseases, may also be detected.

20   In a preferred embodiment, the nucleic acids of the invention find use as probes for toxic bacteria in the screening of water and food samples. For example, samples may be treated to lyse the bacteria to release its nucleic acid, and then probes designed to recognize bacterial strains, including, but not limited to, such pathogenic strains as, Salmonella, Campylobacter, Vibrio cholerae, Leishmania, enterotoxic strains of E. coli, and Legionnaire's disease bacteria. Similarly, bioremediation strategies
25   may be evaluated using the compositions of the invention.

In a further embodiment, the probes are used for forensic "DNA fingerprinting" to match crime-scene DNA against samples taken from victims and suspects.

30   In an additional embodiment, the probes in an array are used for sequencing by hybridization.

The present invention also finds use as a methodology for the detection of mutations or mismatches in target nucleic acid sequences. For example, recent focus has been on the analysis of the relationship between genetic variation and phenotype by making use of polymorphic DNA markers. Previous work
35   utilized short tandem repeats (STRs) as polymorphic positional markers; however, recent focus is on the use of single nucleotide polymorphisms (SNPs). Common SNPs occur at an average frequency of more than 1 per kilobase in human genomic DNA. Some SNPs, particularly those in and around

32

coding sequences, are likely to be the direct cause of therapeutically relevant phenotypic variants. There are a number of well known polymorphisms that cause clinically important phenotypes; for example, the apoE2/3/4 variants are associated with different relative risk of Alzheimer's and other diseases (see Cordor et al., Science 261(1993). Multiplex PCR amplification of SNP loci with

5   subsequent hybridization to oligonucleotide arrays has been shown to be an accurate and reliable method of simultaneously genotyping at least hundreds of SNPs; see Wang et al., Science. 280:1077 (1998); see also Schafer et al., Nature Biotechnology 16:33-39 (1998). The compositions of the present invention may easily be substituted for the arrays of the prior art.

10   In a preferred embodiment, the compositions of the invention are used to screen bioactive agents to find an agent that will bind, and preferably modify the function of, a target molecule. As above, a wide variety of different assay formats may be run, as will be appreciated by those in the art. Generally, the target analyte for which a binding partner is desired is labeled; binding of the target analyte by the bioactive agent results in the recruitment of the label to the bead, with subsequent detection.

15

In a preferred embodiment, the binding of the bioactive agent and the target analyte is specific; that is, the bioactive agent specifically binds to the target analyte. By "specifically bind" herein is meant that the agent binds the analyte, with specificity sufficient to differentiate between the analyte and other components or contaminants of the test sample. However, as will be appreciated by those in the art, it

20   will be possible to detect analytes using binding which is not highly specific; for example, the systems may use different binding ligands, for example an array of different ligands, and detection of any particular analyte is via its "signature" of binding to a panel of binding ligands, similar to the manner in which "electronic noses" work. This finds particular utility in the detection of chemical analytes. The binding should be sufficient to remain bound under the conditions of the assay, including wash steps to

25   remove non-specific binding, although in some embodiments, wash steps are not desired; i.e. for detecting low affinity binding partners. In some embodiments, for example in the detection of certain biomolecules, the dissociation constants of the analyte to the binding ligand will be less than about 10^{-4}-10^{-6} M^{-1}, with less than about 10^{-5} to 10^{-9} M^{-1} being preferred and less than about 10^{-7} -10^{-9} M^{-1} being particularly preferred.

30

Generally, a sample containing a target analyte (whether for detection of the target analyte or screening for binding partners of the target analyte) is added to the array, under conditions suitable for binding of the target analyte to at least one of the bioactive agents, i.e. generally physiological conditions. The presence or absence of the target analyte is then detected. As will be appreciated by

35   those in the art, this may be done in a variety of ways, generally through the use of a change in an optical signal. This change can occur via many different mechanisms. A few examples include the binding of a dye-tagged analyte to the bead, the production of a dye species on or near the beads, the

33

destruction of an existing dye species, a change in the optical signature upon analyte interaction with dye on bead, or any other optically interrogatable event.

In a preferred embodiment, the change in optical signal occurs as a result of the binding of a target analyte that is labeled, either directly or indirectly, with a detectable label, preferably an optical label such as a fluorochrome. Thus, for example, when a proteinaceous target analyte is used, it may be either directly labeled with a fluor, or indirectly, for example through the use of a labeled antibody. Similarly, nucleic acids are easily labeled with fluorochromes, for example during PCR amplification as is known in the art. Alternatively, upon binding of the target sequences, a hybridization indicator may be used as the label. Hybridization indicators preferentially associate with double stranded nucleic acid, usually reversibly. Hybridization indicators include intercalators and minor and/or major groove binding moieties. In a preferred embodiment, intercalators may be used; since intercalation generally only occurs in the presence of double stranded nucleic acid, only in the presence of target hybridization will the label light up. Thus, upon binding of the target analyte to a bioactive agent, there is a new optical signal generated at that site, which then may be detected.

Alternatively, in some cases, as discussed above, the target analyte such as an enzyme generates a species that is either directly or indirectly optically detectable.

Furthermore, in some embodiments, a change in the optical signature may be the basis of the optical signal. For example, the interaction of some chemical target analytes with some fluorescent dyes on the beads may alter the optical signature, thus generating a different optical signal.

As will be appreciated by those in the art, in some embodiments, the presence or absence of the target analyte may be done using changes in other optical or non-optical signals, including, but not limited to, surface enhanced Raman spectroscopy, surface plasmon resonance, radioactivity, etc.

The assays may be run under a variety of experimental conditions, as will be appreciated by those in the art. A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc which may be used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in any order that provides for the requisite binding. Various blocking and washing steps may be utilized as is known in the art.

In a preferred embodiment, two-color competitive hybridization assays are run. These assays can be

34

based on traditional sandwich assays. The beads contain a capture sequence located on one side (upstream or downstream) of the SNP, to capture the target sequence. Two SNP allele-specific probes, each labeled with a different fluorophor, are hybridized to the target sequence. The genotype can be obtained from a ratio of the two signals, with the correct sequence generally exhibiting better

5   binding. This has an advantage in that the target sequence itself need not be labeled. In addition, since the probes are competing, this means that the conditions for binding need not be optimized. Under conditions where a mismatched probe would be stably bound, a matched probe can still displace it. Therefore the competitive assay can provide better discrimination under those conditions. Because many assays are carried out in parallel, conditions cannot be optimzed for every probe

10   simultaneously. Therefore, a competitive assay system can be used to help compensate for non-optimal conditons for mismatch discrimination.

In a preferred embodiment, dideoxynucleotide chain-termination sequencing is done using the compositions of the invention. In this embodiment, a DNA polymerase is used to extend a primer

15   using fluorescently labeled ddNTPs. The 3' end of the primer is located adjacent to the SNP site. In this way, the single base extension is complementary to the sequence at the SNP site. By using four different fluorophors, one for each base, the sequence of the SNP can be deduced by comparing the four base-specific signals. This may be done in several ways. In a first embodiment, the capture probe can be extended; in this approach, the probe must either be synthesized 5'-3' on the bead, or

20   attached at the 5' end, to provide a free 3' end for polymerase extension. Alternatively, a sandwich type assay can be used; in this embodiment, the target is captured on the bead by a probe, then a primer is annealed and extended. Again, in the latter case, the target sequence need not be labeled. In addition, since sandwich assays require two specific interactions, this provides increased specificity which is particularly helpful for the analysis of complex samples.

25

In addition, when the target analyte and the DBL both bind to the agent, it is also possible to do detection of non-labelled target analytes via competition of decoding.

In a preferred embodiment, the methods of the invention are useful in array quality control. Prior to

30   this invention, no methods have been described that provide a positive test of the performance of every probe on every array. Decoding of the array not only provides this test, it also does so by making use of the data generated during the decoding process itself. Therefore, no additional experimental work is required. The invention requires only a set of data analysis algorithms that can be encoded in software.

35

The quality control procedure can identify a wide variety of systematic and random problems in an array. For example, random specks of dust or other contaminants might cause some sensors to give

an incorrect signal-this can be detected during decoding. The omission of one or more agents from multiple arrays can also be detected; that is, since the arrays formulated herein are randomly self-assembled, arrays comprising different statistical distributions of each sensor elements may be made; the invention allows actual QC on the arrays, for "sorting" arrays with good distributions from those

5   with less than ideal distributions. These procedures also allow the determination of sites that do not comprise beads.

An advantage of this quality control procedure is that it can be implemented immediately prior to the assay itself (or after, in some instances), and is a true functional test of each individual sensor.

10   Therefore any problems that might occur between array assembly and actual use can be detected. In applications where a very high level of confidence is required, and/or there is a significant chance of sensor failure during the experimental procedure, decoding and quality control can be conducted both before and after the actual sample analysis.

15   In a preferred embodiment, the arrays can be used to do reagent quality control. In many instances, biological macromolecules are used as reagents and must be quality controlled. For example, large sets of oligonucleotide probes may be provided as reagents. It is typically difficult to perform quality control on large numbers of different biological macromolecules. For example, when large populations of different random oligonucleotides are synthesized, it may be desirable to check that each population

20   is represented. The approach described here can be used to do this by treating the reagents (formulated as the DBLs) as variable instead of the arrays.

In a preferred embodiment, the methods outlined herein are used in array calibration. For many applications, such as mRNA quantitation, it is desirable to have a signal that is a linear response to the

25   concentration of the target analyte, or, alternatively, if non-linear, to determine a relationship between concentration and signal, so that the concentration of the target analyte can be estimated. For example, different probes can show differences in hybridization efficiencies, etc., and it may be desirable to choose appropriate probes under experimental conditions. Thus, for example, when expression profiling is done, and quantitation is desirable, calibration curves can be done to see how

30   individual probes react, and then the probes that give the best response (i.e. linearity at the concentrations and conditions of interest) can be chosen for further assays. Accordingly, the present invention provides methods of creating calibration curves in parallel for multiple beads in an array. The calibration curves can be created under conditions that simulate the complexity of the sample to be analyzed. Each curve can be constructed independently of the others (e.g. for a different range of

35   concentrations), but at the same time as all the other curves for the array.

Thus, in these embodiments, different types of experiments can be done. For example, the sequential

decoding scheme can be implemented with different concentrations being used as the code "labels", rather than different fluorophores. In this way, signal as a response to concentration can be measured for each bead. This calibration can be carried out just prior to array use, so that every probe on every array is individually calibrated as needed. Alternatively, different concentrations of the decoding probe can have different labels.

It should be noted that the assay calibration methods find use in non-random arrays, as well; that is, other types of support-bound nucleic acid arrays can be calibrated using these methods as well. Thus, for example, the sequential addition of different pools of probes, wherein the concentration of the probes is varied, to biochips can allow calibration of any assay system. This type of analysis can also be done on non-random arrays for quality control, to verify the integrity and sequence of the support bound probes, and in assay development to identify good probes.

In a preferred embodiment, the methods of the invention can be used in assay development as well. Thus, for example, the methods allow the identification of good and bad probes; as is understood by those in the art, some probes do not function well because they do not hybridize well, or because they cross-hybridize with more than one sequence. These problems are easily detected during decoding. The ability to rapidly assess probe performance has the potential to greatly reduce the time and expense of assay development. Thus, probes that respond linearly with concentration, show low non-specific binding, or give signals in a particular range, can be chosen for addition to a new array for assays.

Similarly, in a preferred embodiment, the methods of the invention are useful in quantitation in assay development. Major challenges of many assays is the ability to detect differences in analyte concentrations between samples, the ability to quantitate these differences, and to measure absolute concentrations of analytes, all in the presence of a complex mixture of related analytes. An example of this problem is the quantitation of a specific mRNA in the presence of total cellular mRNA. One approach that has been developed as a basis of mRNA quantitation makes use of a multiple match and mismatch probe pairs (Lockhart et al., 1996), hereby incorporated by reference in its entirety. While this approach is simple, it requires relatively large numbers of probes. In this approach, a quantitative response to concentration is obtained by averaging the signals from a set of different probes to the gene or sequence of interest. This is necessary because only some probes respond quantitatively, and it is not possible to predict these probes with certainty. In the absence of prior knowledge, only the average response of an appropriately chosen collection of probes is quantitative. However, in the present invention, that can be applied generally to nucleic acid based assays as well as other assays. In essence, the approach is to identify the probes that respond quantitatively in a particular assay, rather than average them with other probes. This is done using the array calibration

37

scheme outlined above, in which concentration-based codes are used. Advantages of this approach include: fewer probes are needed; the accuracy of the measurement is less dependent on the number of probes used; and that the response of the sensors is known with a high level of certainty, since each and every sequence can be tested in an efficient manner. It is important to note that probes that

5   perform well are selected empirically, which avoids the difficulties and uncertainties of predicting probe performance, particularly in complex sequence mixtures. In contrast, in experiments described to date with ordered arrays, relatively small numbers of sequences are checked by performing quantitative spiking experiments, in which a known mRNA is added to a mixture.

10   In a preferred embodiment, cDNA arrays are made for RNA expression profiling. In this embodiment, individual cDNA clones are amplified (for example, using PCR) from cDNA libraries propagated in a host-vector system. Each amplified DNA is attached to a population of beads. Different populations are mixed together, to create a collection of beads representing the cDNA library. The beads are arrayed, decoded as outlined above, and used in an assay (although as outlined herein, decoding may

15   occur after assay as well). The assay is done using RNA samples (whole cell or mRNA) that are extracted, labeled if necessary, and hybridized to the array. Comparative analysis allows the detection of differences in the expression levels of individual RNAs. Comparison to an appropriate set of calibration standards allows quantification of absolute amounts of RNA.

20   The cDNA array can also be used for mapping, e.g. to map deletions/insertions or copy number changes in the genome, for example from tumors or other tissue samples. This can be done by hybridizing genomic DNA. Instead of cDNAs (or ESTs, etc.), other STS (sequence tagged sites), including random genomic fragments, can also be arrayed for this purpose.

25   All references cited herein are incorporated by reference in their entirety.

CLAIMS

We claim:

1. An array composition comprising:

5   a) a substrate with a surface comprising discrete sites; and

        b) a population of microspheres comprising at least a first and a second subpopulation,

        wherein each subpopulation comprises:

                i) a bioactive agent; and

                ii) an identifier binding ligand that will bind a decoder binding ligand such that the

10   identification of the bioactive agent can be elucidated;

        wherein said microspheres are distributed on said surface.

2. An array composition comprising:

        a) a substrate with a surface comprising discrete sites; and

15   b) a population of microspheres comprising at least a first and a second subpopulation,

        wherein each subpopulation comprises a bioactive agent and does not comprise an optical

        signature, wherein said microspheres are distributed on said surface.

3. A composition according to claim 1 or 2 further comprising at least one decoder binding ligand.

20

4. A composition according to claim 1 or 2 wherein said bioactive agents are nucleic acids.

5. A composition according to claim 1 or 2 wherein said bioactive agents are proteins.

25   6. A method of making a composition comprising:

        a) forming a surface comprising individual sites on a substrate;

        b) distributing microspheres on said surface such that said individual sites contain

        microspheres, wherein said microspheres comprise at least a first and a second

        subpopulation each comprising a bioactive agent and do not comprise an optical signature.

30

    7. A method of making a composition comprising:

        a) forming a surface comprising individual sites on a substrate;

        b) distributing microspheres on said surface such that said individual sites contain

        microspheres, wherein said microspheres comprise at least a first and a second

35   subpopulations each comprising:

                i) a bioactive agent; and

                ii) an identifier binding ligand that will bind at least one decoder binding ligand such

that the identification of the bioactive agent can be elucidated.

8. A method of decoding an array composition comprising
    a) providing an array composition comprising:
        i) a substrate with a surface comprising discrete sites; and
        ii) a population of microspheres comprising at least a first and a second
           subpopulation, wherein each subpopulation comprises a bioactive agent;
       wherein said microspheres are distributed on said surface;
    b) adding a plurality of decoding binding ligands to said array composition to identify the
       location of at least a plurality of the bioactive agents.

9. A method according to claim 8 wherein at least one subpopulation of microspheres comprises an
identifier binding ligand to which a decoding binding ligand can bind.

10. A method according to claim 8 wherein said decoding binding ligands bind to said bioactive
agents.

11. A method according to claim 8 wherein said decoding binding ligands are labeled.

12. A method according to claim 8 wherein the location of each subpopulation is determined.

13. A method of determining the presence of a target analyte in a sample comprising:
    a) contacting said sample with a composition comprising:
        i) a substrate with a surface comprising discrete sites; and
        ii) a population of microspheres comprising at least a first and a second subpopulation
           each comprising a bioactive agent and do not comprise an optical signature;
           wherein said microspheres are distributed on said surface such that said discrete
           sites contain microspheres; and
    b) determining the presence or absence of said target analyte.

14. A method of determining the presence of a target analyte in a sample comprising:
    a) contacting said sample with a composition comprising:
        i) a substrate with a surface comprising discrete sites; and
        ii) a population of microspheres comprising at least a first and a second subpopulation
           each comprising:
               1) a bioactive agent; and
               2) an identifier binding ligand that will bind a decoder binding ligand such that

40

the identification of the bioactive agent can be elucidated;

wherein said microspheres are distributed on said surface such that said discrete

sites contain microspheres; and

b) determining the presence or absence of said target analyte.

5

15. An array composition comprising:

a) a substrate with a surface comprising discrete sites;

b) a population of microspheres distributed on said sites and comprising at least a first and a

second subpopulation, each subpopulation comprising a bioactive agent, wherein said first

:0   subpopulation comprises at least a first optical dye with a first pKa and said second

subpopulation comprises at least a second optical dye with a second pKa, wherein said first

and second pKas are different.

16. A method of making a microsphere array comprising:

15   a) contacting a substrate with a surface comprising discrete sites with a solution comprising a

population of particles; and

b) applying energy to said substrate or said solution, or both, such that at least a

subpopulation of said particles associate onto sites.

20   17. A method according to claim 16 wherein said discrete sites comprise wells.

18. A method according to claim 16 wherein said energy is in the form of agitation.

PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification [6] : **G01N 33/50, 33/543, C12Q 1/68**	**A3**	**(11) International Publication Number:** $\qquad$ **WO 99/67641** **(43) International Publication Date:** $\quad$ 29 December 1999 (29.12.99)

(21) International Application Number: $\quad$ PCT/US99/14387 **(22) International Filing Date:** $\qquad$ 24 June 1999 (24.06.99) **(30) Priority Data:** $\quad$ 60/090,473 $\qquad$ 24 June 1998 (24.06.98) $\qquad$ US $\quad$ 09/189,543 $\qquad$ 10 November 1998 (10.11.98) $\quad$ US **(71) Applicant:** ILLUMINA, INC. [US/US]; Suite 200, 9390 Towne Centre Drive, San Diego, CA 92121 (US). **(72) Inventors:** CHEE, Mark, S.; 155 15th Street #24, Del Mar, CA 92014 (US). STUELPNAGEL, John, R.; 38 Briggs Avenue, Encinitas, CA 92024 (US). CZARNIK, Anthony, W.; 11017 Camino Abrojo, San Diego, CA 92127 (US). **(74) Agents:** BREZNER, David, J. et al.; Flehr Hohbach Test Albritton & Herbert LLP, Suite 3400, 4 Embarcadero Center, San Francisco, CA 94111–4187 (US).	**(81) Designated States:** AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). **Published** $\quad$ *With international search report.* **(88) Date of publication of the international search report:** $\qquad$ 9 March 2000 (09.03.00)

(54) Title: DECODING OF ARRAY SENSORS WITH MICROSPHERES

(57) Abstract

The invention relates to compositions and methods for decoding microsphere array sensors.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER

G01N33/50, G01N33/543, C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC ^{6}

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G01N, C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	WO 99/18434 A (TRUSTEES OF TUFTS COLLEGE) 15 April 1999, claims 1-32.	1-5,8-15
A,P	WO 98/40726 A (TRUSTEES OF TUFTS COLLEGE) 17 September 1998, abstract, claims.	1-18
A,P	WO 98/50782 A (TRUSTEES OF TUFTS COLLEGE) 12 November 1998, claims.	1,4

☐ Further documents are listed in the continuation of box C.   ☐ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 October 1999	21 12. 99

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+ 31-70) 340-3016	SCHNASS e.h.

zum internationalen Recherchen-
bericht über die internationale
Patentanmeldung Nr.

to the International Search
Report to the International Patent
Application Nr.

au rapport de recherche inter-
national relatif à la demande de brevet
international n°

PCT/US 99/14367 SAE 241096

In diesem Anhang sind die Mitglieder
der Patentfamilien der im obenge-
nannten internationalen Recherchenbericht
angeführten Patentdokumente angegeben.
Diese Angaben dienen nur zur Unter-
richtung und erfolgen ohne Gewähr.

This Annex lists the patent family
members relating to the patent documents
cited in the above-mentioned inter-
national search report. The Office is
in no way liable for these particulars
which are given merely for the purpose
of information.

La présente annexe indique les
membres de la famille de brevets
relatifs aux documents de brevets cités
dans le rapport de recherche inter-
national visée ci-dessus. Les reseigne-
ments fournis sont donnés à titre indica-
tif et n'engagent pas la responsibilité
de l'Office.

Im Recherchenbericht angeführtes Patentdokument Patent document cited in search report Document de brevet cité dans le rapport de recherche	Datum der Veröffentlichung Publication date Date de publication	Mitglied(er) der Patentfamilie Patent family member(s) Membre(s) de la famille de brevets	Datum der Veröffentlichung Publication date Date de publication
WO A1 9918430	15-04-1999	AU A1 1349799 / WO C2 9918430	27-04-1999 / 06-05-1999
WO A1 9840724	17-09-1998	AU A1 6468798	29-09-1998
WO A2 9850782	12-11-1998	AU A1 7357798 / WO A1 9850782	27-11-1998 / 12-11-1998